

INTERACTIVE MACHINE LEARNING SYSTEM FOR AUTOMATED ANNOTATION OF INFORMATION IN TEXT

5

ABSTRACT OF THE DISCLOSURE

10 An interactive machine learning based system that incrementally learns, on the basis of text data, how to annotate new text data. The system and method starts with partially annotated training data or alternatively unannotated training data and a set of examples of what is to be learned. Through iterative interactive training sessions with a user the system trains annotators, and these are in turn used to discover more annotations in the text data. Once all of the text data or a sufficient amount of the text data is annotated, at the user's discretion, the system learns a final annotator or
15 annotators, which are exported and available to annotate new textual data. As the iterative training process occurs the user is selectively presented for review and appropriate action, system-determined representations of the annotation instances and provided a convenient and efficient interface so that context of use can be verified if necessary in order to evaluate the annotations and correct them, where required. At
20 the user's discretion, annotations that receive a high confidence level can be automatically accepted and those with low confidence levels can be automatically rejected.

\\COM\192177.1